

Product Analysis:

Analyze products, their popularity, ranking, rating and price elasticity effect.

Dataset preparation:

```
[10]:
```

| | order_id | customer_id | order_date | total_amount | status | product_id | quantity | product_name | category | price | price_logged |
|---|----------|-------------|------------|--------------|-----------|------------|----------|----------------|-------------|-------|--------------|
| 0 | 1 | 101 | 2023-01-05 | 250 | completed | 1 | 2 | Wireless Mouse | Electronics | 25 | 3.258097 |
| 1 | 1 | 101 | 2023-01-05 | 250 | completed | 3 | 1 | Phone Case | Accessories | 15 | 2.772589 |
| 2 | 2 | 102 | 2023-01-07 | 150 | cancelled | 4 | 1 | Headphones | Electronics | 100 | 4.615121 |
| 3 | 3 | 103 | 2023-01-10 | 300 | completed | 6 | 1 | Smartwatch | Electronics | 150 | 5.017280 |
| 4 | 3 | 103 | 2023-01-10 | 300 | completed | 8 | 1 | Laptop | Electronics | 1000 | 6.908755 |

Dataframe is derived by combining multiple dataframes such as orders, orders items and products, based on common fields product_id and order_id.

This gives complete detail of product analysis.

Here the extra column is included in dataframe, price_logged. This is due to the wide range of product price. To accumulate with various functions and algorithms. This standardization is done during previous steps of data preparation and standardization.

Ranking:

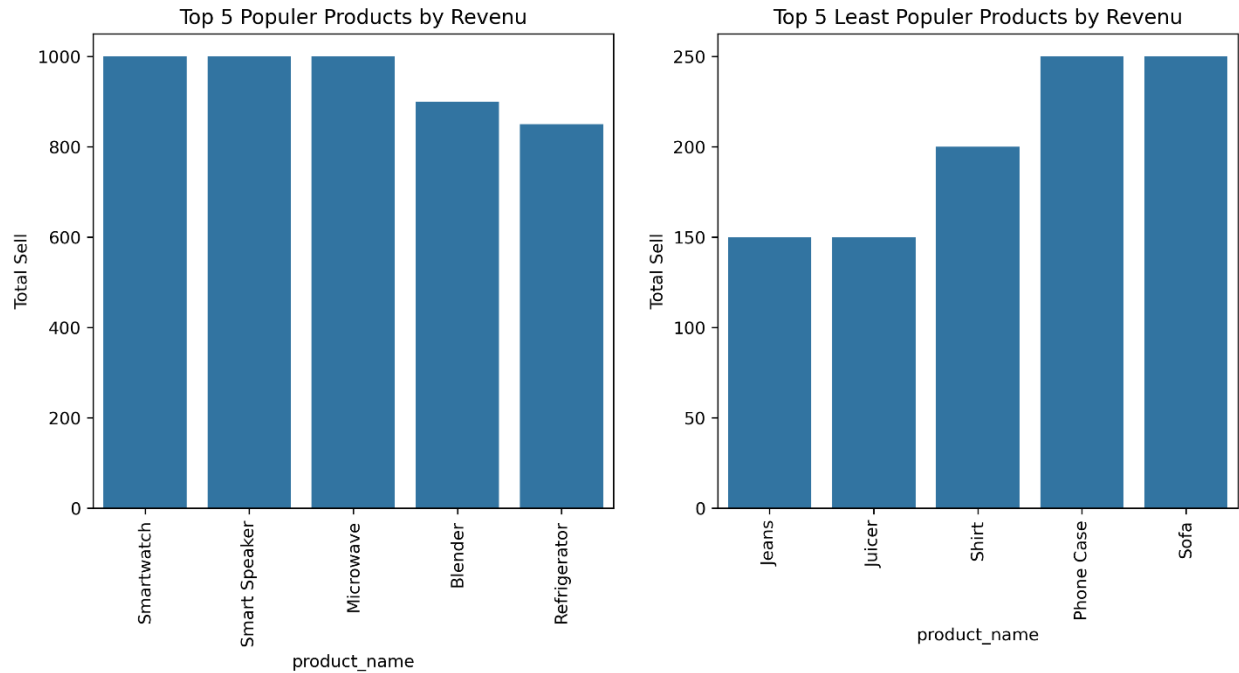
The ranking is based on:

- Order frequency
- Total sell

| | Total Sell | Purchase Frequency |
|-----------------|------------|--------------------|
| product_name | | |
| Air Conditioner | 550 | 1 |
| Blender | 900 | 1 |
| Chair | 850 | 2 |
| Coffee Maker | 250 | 1 |
| Desk Lamp | 320 | 2 |

Top 5 products by revenue:

Bottom 5 products by revenue:



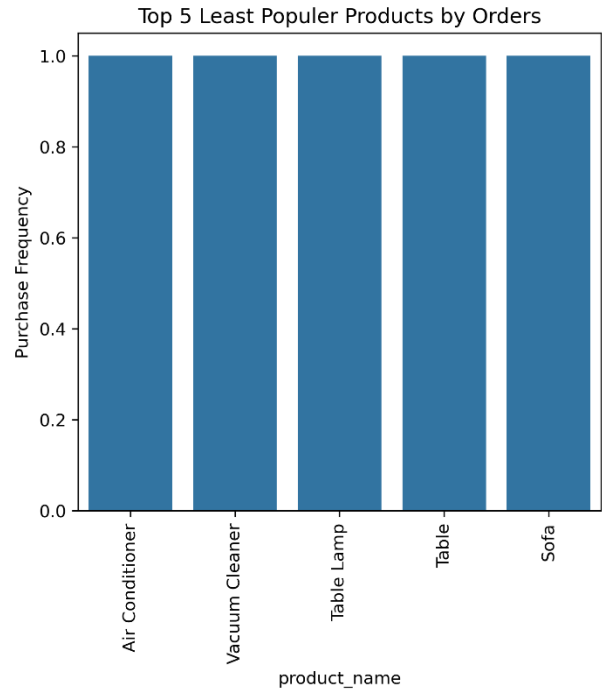
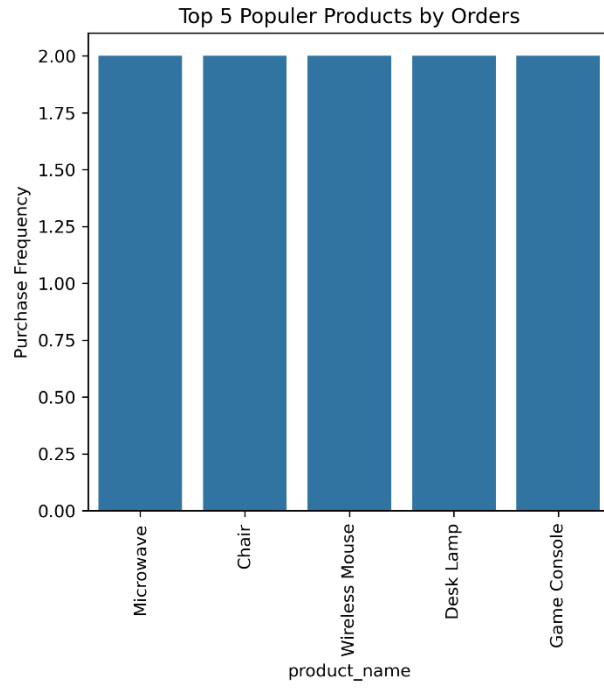
Top selling products are the key contributor to revenue whereas least sold products are the one who needs extra attention.

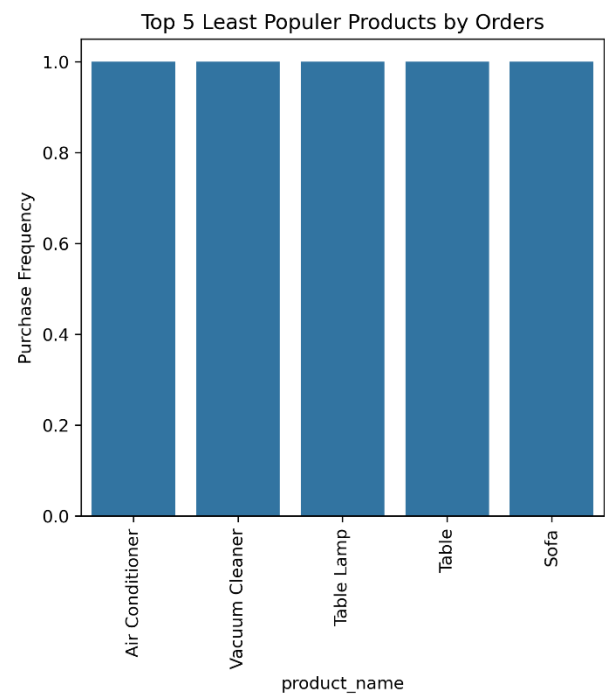
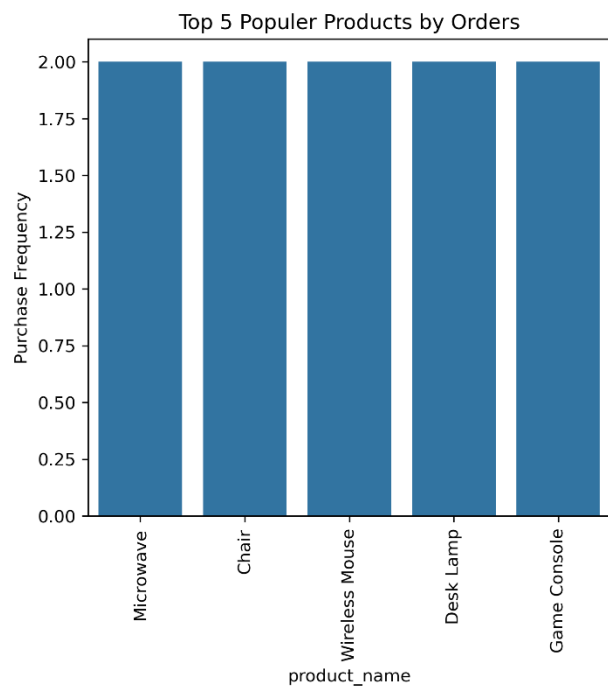
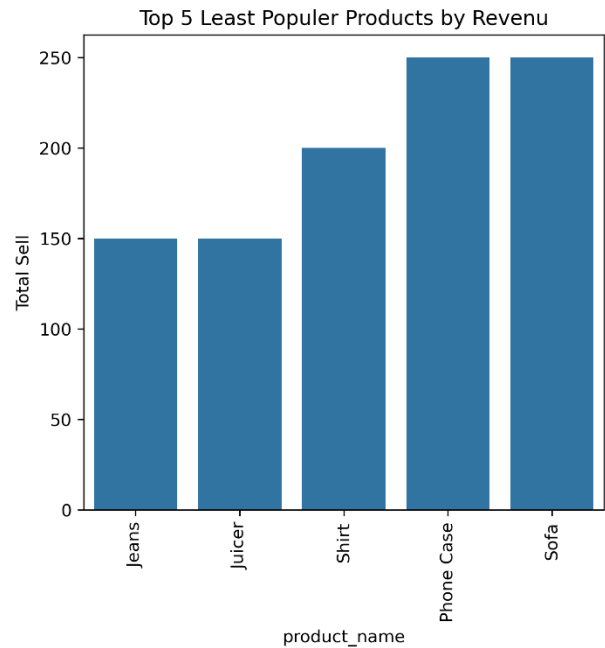
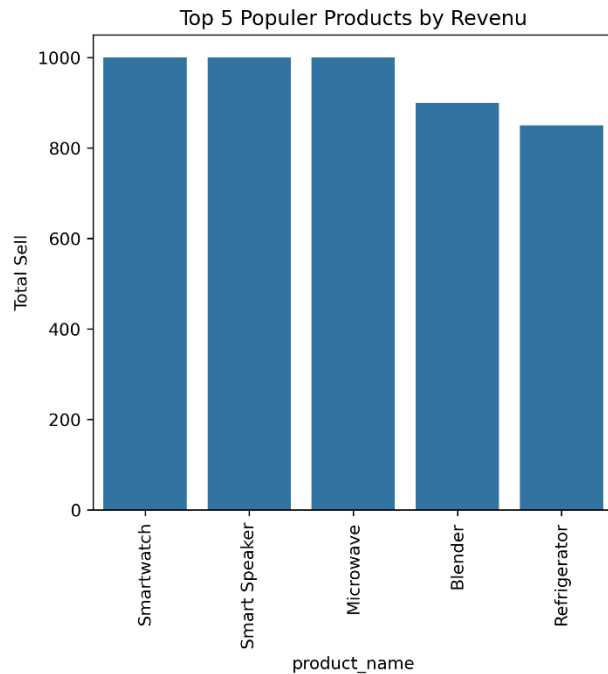
These include further analysis such as:

- These are lest revenue generators because of their price
- Or they have less order frequency
- Or they have low ratings

Top 5 products by order frequency:

Bottom 5 products by order frequency:

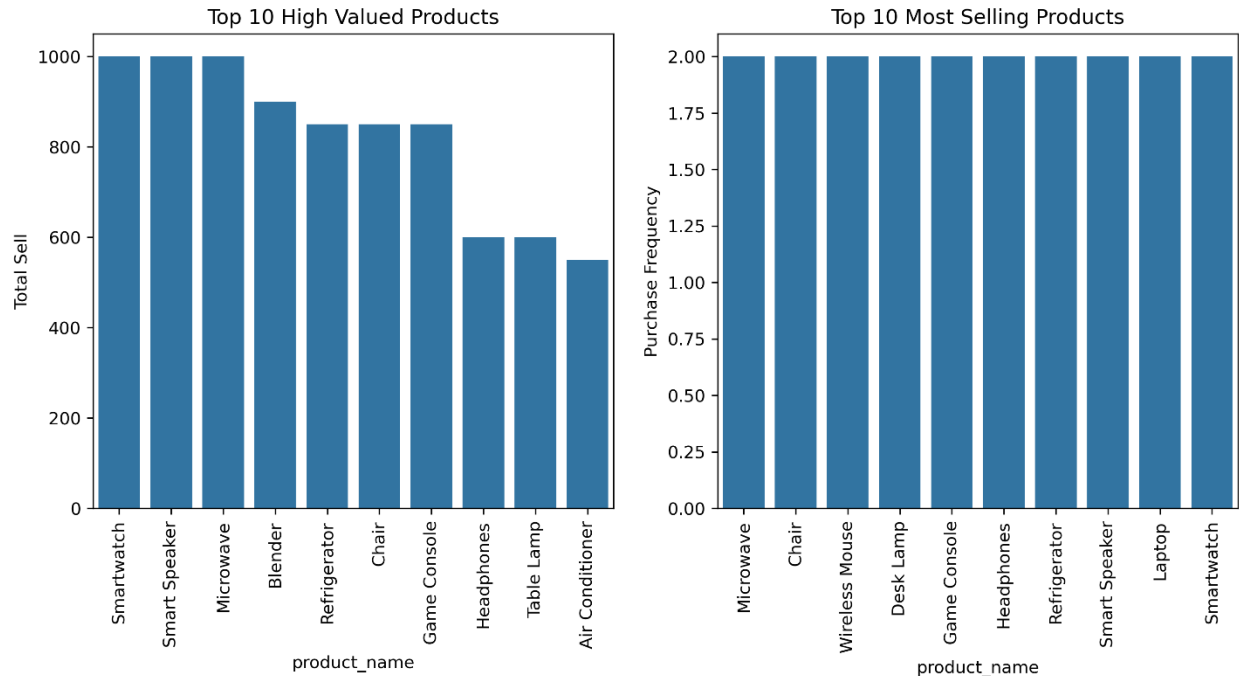




The above dashboard gives different dimensions to analysis. It reveals further hidden patterns and trends in revenue generation.

Top revenue generators such as **smartwatches, smart speakers** and least generators such as **jeans, jeans, juicer** are not one who is brought frequently. The most optimal

products are the one who generates revenue along with being top rank in order frequency, such as **microwave**.



Further expanding the range, it reveals more potential products such as

- Microwave
- Gaming consol
- Headphones

Price Elasticity Analysis:

This is kind of analysis done on product price to check effect of product price change on product sale:

What if:

Price increased:

An increase in price will divert customers to competitors.

Price dropped:

Dropping price will bring more customers with some part of competitors. But this will affect revenue generation due to less profit margin

The sweet spot is the price that gives good profit margin and revenue generation, keeping price optimum.

However, this analysis needs price history data, the record that tells how much price variation happened in the past and how it affected the sales.

In current data available, there is no price history available, the PEA cannot perform.

Alternative Analysis:

Price vs sales:

Focus is to check relation between price and sale.

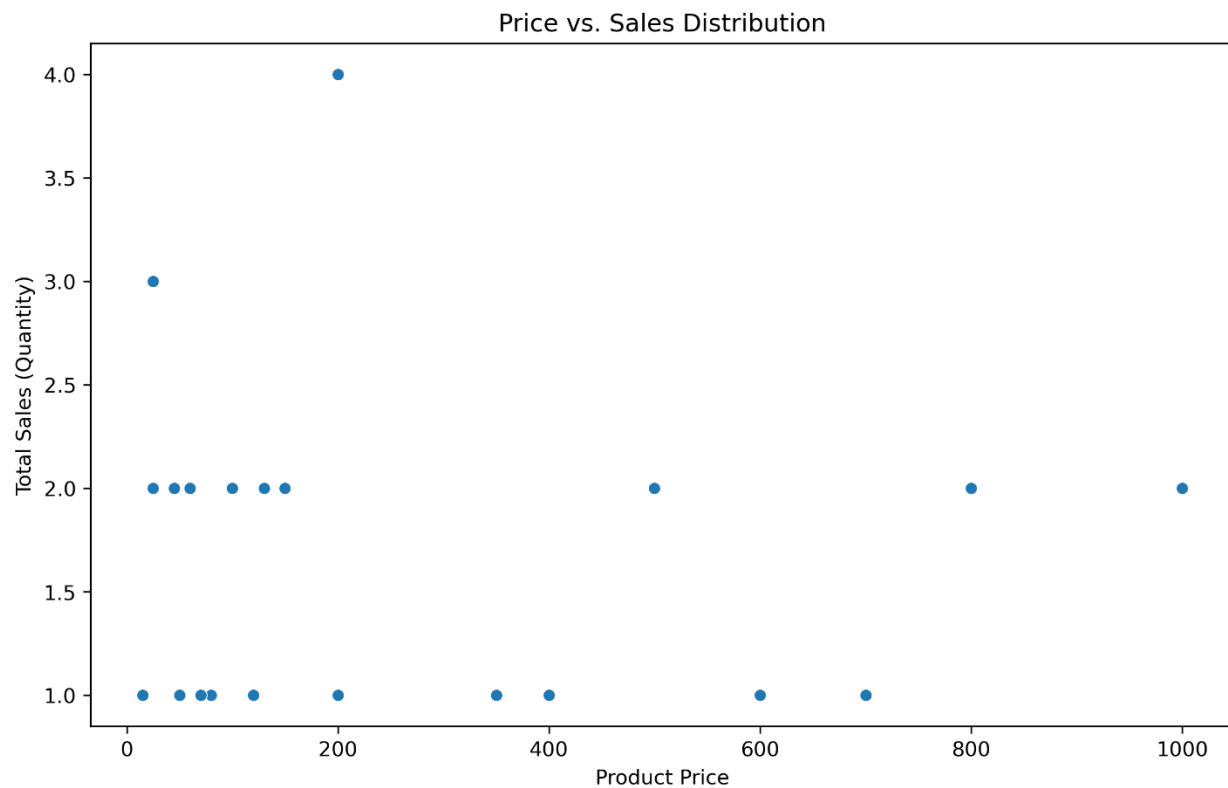
Based on relation, the sale is categories as

Price sensitive:

This tells more about customer group and their purchase mindset, more quantity sold with low price, this means that customer group is price sensitive

Premium product based:

more order of high priced, tells customers less concern about price and selling costly product is easy, hence generating more revenue



Data points are clustered around:

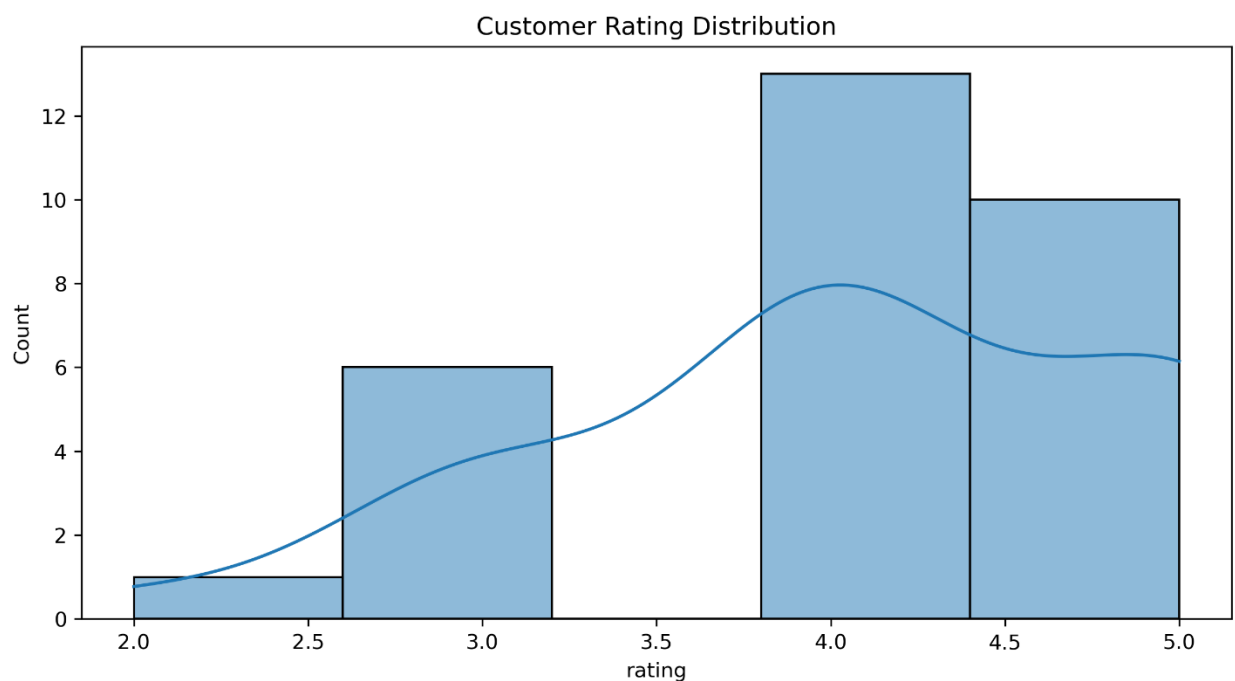
Price range = 0 to 600

Order count = 1-2

This indicates most of the revenue generated is by low – to -medium priced products bought frequently.

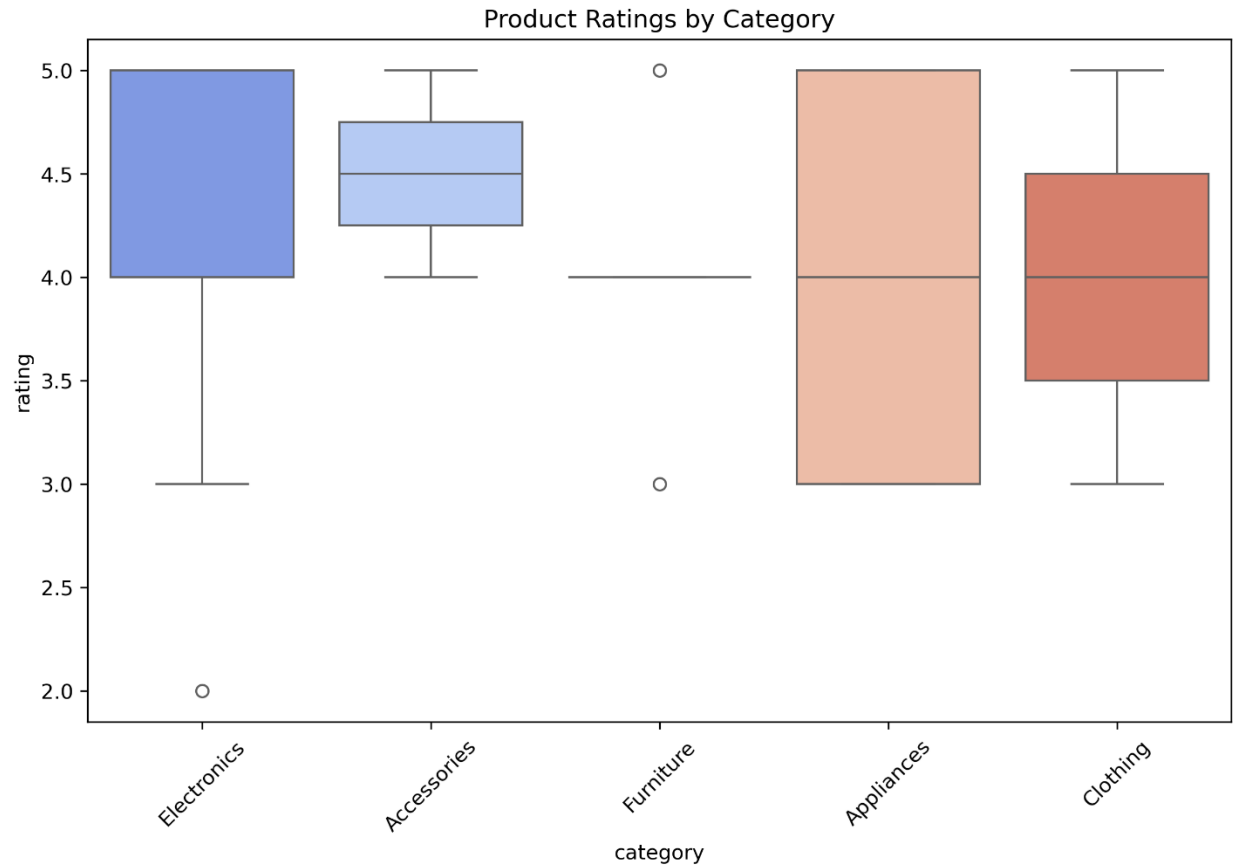
Product rating and sales:

Product rating may also play a role in sales. Low rating reflects negative feedback, which indirectly affects the customer's decision whether to buy or not.



Distribution chart: shows the data is clustered on right side, this explains the most of products are positively rated.

Further analysis includes, category wise rating check, to identify if any category is performing low due to lower rating.



Electronics: range = 4 to 5

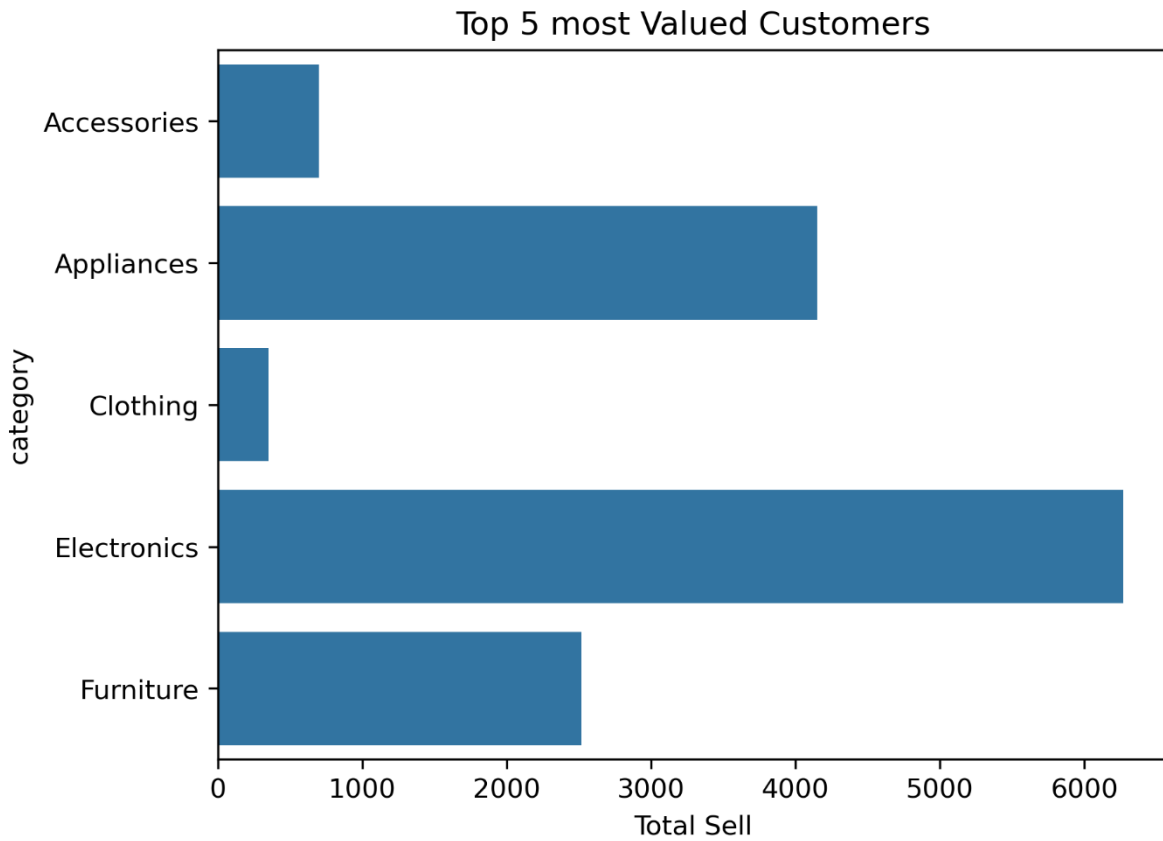
Accessories: range = 4.25 to 4.75

furniture: range = 4

appliances: range = 3 to 5

clothing: range = 3.75 to 4.75

Looking at above figures, the category appliance is having wide range of rating varying from 3 to 5. this raises concern about sales contribution by Category **Appliances**.



This indicates that variation in rating doesn't affect the sales.

Corelating rating with selling price:

To verify if the company has priced products optimal or they are overpriced.



| | rating | total_amount |
|--------------|----------|--------------|
| rating | 1.000000 | 0.242451 |
| total_amount | 0.242451 | 1.000000 |

0.24 correlation indicates that price barely affects the rating of product.

Rating distribution vs sales:

Clustered datapoints around the top left corner indicate most of revenue generated by top rated products. And this concludes that rating doesn't that much affect the sales.

Rating = 2: There's only one data point at a rating of 2, with relatively low sales. This could be an outlier or represent a specific product.

Rating = 3: There's a wide range of sales at a rating of 3, from low to relatively high.

Rating = 4: Similar to rating 3, there's a spread of sales values.

Rating = 5: Again, there's a wide range of sales. Some of the highest sales values are observed at this rating, but there are also data points with lower sales.

